DOCUMENT RESUME

ED 175 895 TM 008 735

AUTHOR Rabinowitz, Stanley N.; Pruzek, Robert
TITLE A Modern Perspective on Some Common Factor Methods: Reinstatement and Rehabilitation of Old Methods for New Purposes.
PUB DATE 78
NOTE Paper presented at the Annual Meeting of the Northeastern Educational Research Association (Ellenville, New York, October 25-27, 1978); Best copy available

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Bayesian Statistics; *Comparative Statistics; Educational Research; *Factor Analysis; *Mathematical Models; Matrices; *Statistical Analysis; Statistical Studies; Tests of Significance
IDENTIFIERS Maximum Likelihood Estimation; *Principal Components Analysis; Whittle (P)

ABSTRACT

Despite advances in common factor analysis, a review of 89 studies published in four selected journals between 1963 and 1976 indicated that behavioral scientists preferred principal components analysis, followed by varimax or orthogonal rotation. Resultant row sums of squares of factor matrices from principal component analyses of real data sets had a relatively low correlation with similar matrices from the most widely used common factor methods. A slight generalization of Whittle's method of common factor analysis was shown to provide a number of distinctive advantages over component analyses and Joreskog's likelihood methods. Finally, guidelines were provided to assist prospective factor analytic users in deciding when their problems might best be approached by common factor methods, and when other multivariate structural methods might be more effective or at least complementary. A Bayesian approach to these problems was advocated in this final section. (Author/CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

- ED175895

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A MODERN PERSPECTIVE ON SOME COMMON FACTOR METHODS REINSTATEMENT AND REHABILITATION OF OLD METHODS FOR NEW PURPOSES

STANLEY N. RABINOWITZ

AND

ROBERT PRUZEK
STATE UNIVERSITY OF NEW YORK AT ALBANY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
S. Rabinowitz

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM008 735

Paper presented at the annual meeting of the Northeastern Educational Research Association, October 25-27, 1978, Ellenville, NY.

As a topic for research in psychometrics or multivariate analysis, no method has received more attention than common factor analysis (CFA). Factor methods have been under development for almost three-quarters of a century, and several comprehensive books are now available from which to learn about various CFA methods and their interrelationships. Not only has the subject of factor analysis been intrinsically fascinating to large numbers of theoreticians, but the problems with which factor analysis are designed to deal, continue to appear frequently in the behavioral sciences.

Yet, despite the relatively advanced stature of factor methodology, behavioral scientists have come to use common factor methods only infrequently, opting most often for component approaches to structural model analysis. In this paper, we have three main purposes. First, we survey five educational and psychological journals for the years 1975-1976, following by one decade the survey of Glass and Taylor (1966). This examination of published research serves especially to document the prevalence of component analytic versus factor analytic methods in applied behavioral research. Next, we examine certain relatively old, and obviously forgotten, forms of common factor analysis which appear to us to merit careful consideration for modern applied CFA research. These methods have especially close ties with principal component (PC) methods and seem to us to provide a basis for a potentially more meaningful interpretation of conventional PCA results and extension beyond these results. Finally, we present a point of view concerning how some of these older CFA methods may reasonably be expected to serve work-a-day needs in factor applications. The issue of confirmatory versus exploratory methods is discussed from a Bayesian perspective.

Glass and Taylor (1966) surveyed four educational and psychological research journals to determine which methods of factor analysis were the most widely used by researchers in these fields. Four journals--American Educational Research Journal (1965-March, 1966), Journal of Educational Measurement (1964-1965), Journal of Educational Psychology (1963-April, 1966), and Educational and Psychological Measurement (1963-Spring, 1966)--were examined. Of the 46 factor analyses studies they located, 10 gave too little information to permit classification.

Of the 36 studies which could be classified, 23 were found to be principal component analysis, beginning from a matrix of product-moment correlations. In 16 of these 23 cases, varimax rotation was used. Only four cases employed non-orthogonal factor transformations.

In order to examine the present state of factor analytic practice ten years after the survey of Glass and Taylor, we have reviewed the same four journals for the years 1975-1976.

Of a total of 53 factor analyses found for this two-year period, 33 employed the principal component solution, nearly all of which followed with varimax or some unspecified "orthogonal" rotation. Thirteen studies provided too little information to permit any classification, however, we are not wholly sure of the meaning in four cases in which "principal factor solutions" and "iterative principal components" methods were mentioned. Of the 40 classifiable studies, only five used a non-orthogonal transformation.

As in 1966, the criterion for choosing the number ~~of factors~~ of factors to extract was almost exclusively the number of eigenvalues greater than unity. Half of those studies which used Principal Component Analysis reported using this technique. That the number of eigenvalues greater than unity is the default option for the most popular canned factor analysis programs (SPSS and BMDP) is perhpas the strongest reason for the popularity of this criterion.

Four other studies reported choosing the number of factors by a more complex examination of the eigenvalues than the one above, while three used Cattell's Scree Test to decide. It seems significant that none reported chi square tests to determine the statistical adequacy of their solutions (see Joreskog, 1967, for example).

Finally, it is also a curious finding that none of the studies we surveyed reported more than a single method of factor "extraction" whereas in only four cases were alternate transformation methods compared. It would appear from these findings that there have been few advances, or changes of any real significance, in the period since Glass and Taylor's (1966) review. We might add that while several fine technical monographs and textbooks have been published on the topic of factor analysis since 1966, most notably Mulaik (1972), the more general methodological texts in psychology seem in virtually all cases to be devoid of clear and revealing discussions of modern factor methods. In particular, a student cannot expect to find an informative discussion concerning the distinctions among major methods of component and factor analysis. As our results in the next section indicate, different methods do, indeed, tend to yield results in practice which are incompatible with one another.

As noted earlier, our second purpose in this paper is to reconsider certain forms of common factor analysis (CFA) which have particularly close ties to principal component analysis. The CFA methods to which we refer have been discussed most fully by Whittle (1952). However, they have also been considered by Bartlett (1950, 1951) and Lawley (1953). It is rather easy to provide the essential information on Whittle's methods, but the reader might do well to review the basic model for common factor analysis as discussed for example in Chapter 5 of Mulaik (1972).

The usual assumptions of CFA lead to the equation:

$$\Sigma = FF' + U^2 \qquad (1)$$

where Σ is a pxp variance-covariance matrix for p random variables, F is a pxm matrix of common factor coefficients (for orthogonal factors, here) and U^2 is a pxp diagonal matrix of uniqueness factor variances. Although the CFA model is often written in terms of a pxp correlation matrix, $R = (\text{Diag})\Sigma(\text{Diag})$, it is important in principle to distinguish between CFA analyses of Σ and of R.

Given the CFA model, and a sample covariance (or correlation) to estimate Σ (or R), the typical approach to analysis is to estimate F and U^2, conditional on some choice for m, the number of common factors.

In fact, however, there are serious indeterminacy problems connected with the simultaneous estimation of F and U^2 in general CFA applications. (We shall mention some of the specific problems later.)

The Whittle method can be seen to avoid many of these problems in a way which seems to carry implications for a serious reexamination of principal component methods and perhaps CFA more generally, as well. Of particular interest, in the context of this paper, is the light which a detailed examination of the Whittle method sheds on PCA methods. In view of the obvious heavy use that PCA methods have been getting, it seems important to understand them as fully as possible.

The Whittle Method:

Whittle's (1953) idea was to consider the models

$$\Sigma = FF' + \phi^2 I \qquad (2a)$$

and $$R = FF' + \phi^2 I \qquad (2b)$$

in place of (1). Thus, he assumed that uniqueness variances were all equal to one another with respect to the original scaling of the observed variables, as in (2a), or proportional to the original (population) variances, as in (2b). He called these cases of constant and proportional error variance the additive and multiplicative cases, respectively. Model (2a), in Whittle's hands, calls for a direct eigenanalysis of the observed sample covariance matrix, whereas model (2b) calls for an eigenanalysis of the pxp correlation matrix. The additive model is more tractable for inquiries about statistical distributions, whereas the multiplicative model relates more directly to the typical practice of principal component analysis, starting from a matrix of product-moment correlations.

Using Whittle's methods, the matrix F of sample factor coefficients is simply computed as

$$F_W = Q\,(D_\lambda^2 - \phi^2 I)^{\frac{1}{2}} = \phi \cdot Q\,(\phi^{-2} D_\lambda^2 - I)^{\frac{1}{2}} \qquad (3)$$

where Q and D_λ^2 are simply eigenvectors and values of $\hat{\Sigma}$ or $\hat{R}$. Of course, in practice, only the largest eigenvalues and corresponding vectors are used in forming F_W.

The quantity ϕ^2 may either be prespecified or estimated from the data. There is good reason in most circumstances to define ϕ^2 as the average value of the p-m smallest eigenvalues of the covariance or correlation matrix (cf. Joreskog, 1962). Thus,

$$\phi^2 = \frac{1}{p-m} \sum_{j=m+1}^{p} \lambda_j^2$$

defines the usual estimator of the "common" uniqueness variance. It should be clear that the Whittle F matrix is, for either (2a) or (2b), formally very similar to the F for principal component analysis:

$$F_{PCA} = Q D_\lambda$$

Again, in practice, the typical approach is to use incomplete component analysis--most frequently the number of eigenvalues greater than one if a correlation matrix is analyzed.

It appears that an especially effective way to characterize the similarities and differences between PCA and Whittle solutions is through the use of numerical examples. Our first example uses a simple population matrix. It was inspired by both Bartlett (1950) and Whittle (1952). Let R_{3x3} be a population correlation matrix in which $\rho_{12} = .64$, and $\rho_{13} = \rho_{23} = 0$. A direct eigenanalysis (PCA) of this R yields the eigenvalues $\lambda_1^2 = 1.64$, $\lambda_2^2 = 1.0$ and $\lambda_3^2 = .36$. The corresponding (unit-length) eigenvectors are $\underset{\sim}{q}_1^1 = \angle .707, .707, 0 \rfloor$, $\underset{\sim}{q}_2^1 = \angle 0., 0., 1.0 \rfloor$ and $\underset{\sim}{q}_3^1 = \angle .707, -.707, 0 \rfloor$. Thus, a single component, the one corresponding to the largest eigenvalue, yields a single component $F_{PCA}^1 = \angle \underset{\sim}{q}_1^1 \rfloor \lambda_1 = \angle .906, .906, 0 \rfloor$, showing that the value ρ_{12} is not accurately reproduced using a single component, FF^1. However, if ϕ^2 is set at .36, then the one factor Whittle $F_w^1 = \angle \underset{\sim}{q}_1^1 \rfloor (\lambda_1^2 - \phi^2)^{\frac{1}{2}} = \angle .8 \ .8 \ 0 \rfloor$, showing that $F_w F_w^1$ reproduces exactly the off-diagonal elements in R, viz., $\rho_{12} = .64$. Table 1 neatly summarizes the results of this analysis. However trite this example may be, it clearly shows the shortcomings of PCA, particularly in light of the success of the Whittle procedure.

The use of PCA involves the implicit or tacit assumption that all observations in the original (Nxp) <u>data</u> matrix are observed <u>without error</u>, whereas CFA involves the fundamentally different assumption that each of the original observations is observed <u>with</u> error.

As the above example illustrates, parsimony in analysis (use of a single component here) is generally not compatible with the assumption that all variables are observed without error.

Moreover, it is difficult (at least for us) to imagine when or how the no-error assumption could be justified in applied studies, at least in the behavioral sciences. What virtues there are for PCA methods appear to be associated with the extent to which they approximate more realistic multivariate models for behavioral applications.

As it is ordinarily applied, CFA, in general, involves the further assumption that for any variable the error (or uniqueness) is properly defined on the basis of that variable's covariance with the (p-1) others in the same set or battery. The Whittle method does not necessarily involve this latter assumption, however. This is because with Whittle method, error (or uniqueness) variances are taken as "known" in advance, possibly, indeed probably, on the basis of information which goes beyond consideration of the other observed variables. A good example would be the use of variable reliability information to prespecify uniqueness variances.

A slight generalization of the Whittle method, although it was actually implied by Whittle, is to specify the model

$$\Sigma \text{ or } R = FF^1 + \phi^2 D_u^2 \qquad (2c)$$

where D_u^2 consists of uniqueness variances that are defined in advance of analysis, up to the constant of proportionality ϕ^2. It should be clear that this model is fundamentally less restrictive than (2a) or (2b). The different nonzero entries in D_u^2 might reasonably be set on the basis of known (or perhaps, guessed) reliabilities of the individual variables. Pruzek (1977) discussed this generalized model in some detail.

Table 2 exhibits three sets of factor results for a 12x12 matrix of correlations taken from Hauser (1973) for which n = 3427. Orthogonally transformed (equamax) F matrices (all 12x5) are given for principal component analysis, Whittle CFA (model 2b) and maximum likelihood CFA. We provide the uniqueness variance estimates corresponding to the respective F matrices, the complements of the row sums of squares of F in each case. U^2 estimates are examined rather than commun ality estimates because the former are of more direct interest here.

These matrices of factor coefficients and their corresponding uniqueness values appear to suggest at least moderately incompatible interpretations for the different methods. Curiously, the Whittle uniqueness variance estimates correlate much more highly with the maximum likelihood $\hat{u}^2$ values than do the PCA "uniqueness estimates." This, despite the fact that the Whittle uniqueness variance estimates are even more restricted in range than the PCA values. Certain other descriptive statistics relating to the adequacy of the solution are also provided.

Table 3 provides information pertaining to a comparison of five different uniqueness estimators for six well-known real data sets problems. The data sets include Holzinger (Harmon, 1960), (n = 145, p = 24, m = 4,5); Emmett (Joreskog, 1967) (n = 211, p = 9, m = 3); Schneider (Mulaik, 1975) (n = 270, p = 12, m = 4,5); Hilton (Joreskog, 1973) (n = 720, p = 10, m = 3,4); Hauser (1973) (n = 3427, p = 12, m = 4,5), and the 8 Physical Variable Set (Harmon, 1960) (n = 305, p= 8, m = 3). For four of the data sets, results are presented for two different numbers of factors, resulting in ten solutions.

Initially, means and standard deviations are presented for the uniqueness variance estimators. Maximum likelihood $\hat{u}^2$ values are presented first to provide a comparative standard because at least for moderate-to-large sample sizes, they are generally accepted as the theoretically most appropriate estimators. Jöreskog (1967) is generally credited with the first acceptable operational solution for the maximum likelihood CFA estimators. Also, Jöreskog studies several of these same data sets.

The first, third and fifth type of u^2 estimators were the initial focus of this study; these are the maximum likelihood, the Whittle (model 2b) and the PCA values, respectively. More will be said about the second type of estimator later. The fourth u^2 estimators ($1-r^2$) are the frequently used complements of the squared multiple correlations.

Note first that for the first three uniqueness estimators, the average values are very similar to one another for each solution. Yet, the S^2 values are larger, while the PCA $\hat{u}^2$ values are smaller than the initial values. The bias of both of the latter estimators is easy to demonstrate analytically; nevertheless, the magnitudes of bias are hard to assess without empirical data such as these.

Note that the maximum likelihood estimators are in all cases more highly variable than the Whittle or PCA estimators. This, too, is predicted by theory, at least for the Whittle model (2c).

The results of primary interest in Table 3 are the coefficients of correlation between the various estimators. It is clear that in eight of ten cases the PCA estimators correlate distinctly higher with the maximum likelihood $\hat{u}^2$ values, in relation to the preceeding estimators. In only one case do the PCA $\hat{u}^2$ values correlate better with the maximum likelihood values than the Whittle values.

These results seem to suggest that the relatively minor modification of the PCA F matrix which is associated with using the Whittle method may routinely provide a substantial improvement in terms of the goal of trying to recover uniqueness information for the different variables in a battery. Indeed, it seems disconcerting in light of their prevalence, that the PCA methods appear to provide so little information about the relative error variances in real data problems like these. In five of ten cases, the correlation between Whittle and maximum likelihood $\hat{u}^2$ values exceeds .851, whereas only one PCA $\hat{u}^2$ set passes this test. One set of PCA $\hat{u}^2$ values even correlates negatively with the maximum likelihood $\hat{u}^2$ values.

The second set of $\hat{u}^2$ estimators, labeled as "MLFA/2 iter" were generated by starting the maximum likelihood iteration procedure from the Whittle $\hat{u}^2$ values, going through two cycles of what is usually called Thomson's algorithm, to find how much the Whittle values move toward the maximum likelihood values in two cycles and how variable they have become. The essential reasoning comes from the work of Stein (James and Stein, 1961) and involves the general proof that in multiparameter estimation there is good reason to shrink, say, least squares estimators toward their

common mean. We may thus regard the Whittle $\hat{u}^2$ estimators as "shrunken" maximum likelihood estimators, although, perhaps shrunken "too much." The "MLFA/2iter" estimators can be seen as providing an intermediate degree of shrinkage, with special virtues, perhaps, when sample sizes are moderate-to-small. This idea also can be tied in to Bayesian estimation, where some form of vague knowledge is introduced about the variability of the uniqueness estimators. At any rate, these estimators, however arbitrary, appear to have the characteristics we desired for them.

The final data in Table 3 are the set of chi square statistics for the maximum likelihood and the Whittle procedures for these data sets. We include the data only to draw attention to the possibility of making a statistical assessment of the adequacy of Whittle methods using Bartlett's (1950, 1951) theory. For most of these data sets, the Whittle chi square statistics are so much larger than their maximum likelihood analogues that more factors might appear to be indicated. It is well-known, however, that common factors may account for "significant" amounts of variance when using a conventional alpha value even though these factors seem to be of little practical importance. Perhaps the more serious problem which sound CFA methods may help to avoid is the interpretation of more factors than seem to be indicated statistically.

Our final purpose is to begin to express a point of view about structural analytic studies which provides a framework for the preceding studies of CFA methods. What we have to say here is only suggestive of a much more lengthy exposition which has yet to be developed. Our review of orthodox practice in this field suggests that some fundamentally new conceptualizations are rather badly needed at this time.

As Bayesian's are fond of pointing out classical statistical methods typically do not permit a formal introduction of relevant prior beliefs in analysis. That is, classical methods do not permit operational hypothetico-deductive use of Bayes' theorem. But, a little reflection tells us that particularly in multivariate studies, even formal Bayesian methods cannot, in general, accommodate <u>all</u> relevant prior information. We are left with a problem in any particular applied study of using a "tractable" method, which is "accommodated to" (or is not fundamentally inconsistent with) our prior beliefs in our particular circumstances. Classical multivariate methods appear, therefore, to be essential to serve our applied needs--at least for the recognizable future

We believe that the key point is to choose methods which are "essentially consistent" with our held beliefs, however vague, whether using classical, Bayesian, empirical Bayesian, or any other methods.

The past decade has seen the development of several general methods of accommodating highly specific, as well as less specific, forms of covariance structure analysis. Joreskog (1970) and McDonald (1974) have been especially instrumental in developing the possibilities. These are of course broad developments in classical statistical (and also psychometric) methods. It is now obvious that CFA methods such as those we have focused on can be seen as special cases within this covariance structure framework. In particular, the Whittle methods can be seen as methods which accommodate vague or loosely held prior beliefs about structural relations, such as a belief in "simple structure" for the factor coefficients matrix, while at the same time permitting one to incorporate particular prior beliefs about "errors." The most general of the Whittle models (model 2c) is, to us, especially attractive for a wide range of possible applications. One need only prespecify relative error variances and the number of factors for a given analysis to proceed.

Advantages of the Whittle CFA methodology include the following: In relation to CFA methods, p-1 fewer parameters are estimated so that there is reason to expect greater sensitivity than conventional CFA methods, at least when one's prior uniqueness values are "realistic." Whittle methodology also avoids entirely the common "Heywood" problems that occur for maximum likelihood CFA methods (when uniqueness variance estimates are "nonpositive"). The Whittle procedures are computationally just as efficient as principal component methods. Thus, they are

very much more efficient than maximum likelihood methods. A further virtue is that they are easy to program and to interpret, no changes being required in conventional transformation algorithms. Finally, Whittle methods can be employed with statistical tests for the number of common factors, so that CFA becomes a truly falsifiable model in contrast with purely data reduction methods like principal component analysis.

Especially important is the fact that the use of Whittle methods should tend to encourage careful attention to the design of batteries of observed variables. Thus, factorial redundancy (low rank), and "simple structure," become important design principles which serve not only to accommodate, but also to help shape, prior beliefs about structure, particularly at the beginning stages of research in given domains. Sequentially articulated studies, using the results of one study to help focus the succeeding ones, seem far more likely to emerge when using methodology which provides for realistic beliefs about errors, and which provides "posterior" estimators of error variances which, as we have shown, tend to correlate so highly with maximum likelihood estimators. It is especially difficult to justify conventional PC methodology in the framework of an inductive system that involves the concept of a priori beliefs in an explicit way. Nevertheless, operationally, it is not demanding very much to ask investigators to modify their PC algorithms to implement the Whittle methodology.

References

Bartlett, M.S., Tests of significance in factor analysis. British Journal of Psychology, 1950, 3, 77-85.

Bartlett, M.S. A further note on tests of significance in factor analysis, British Journal of Psychology, Statistical Section, 1951, 4, 1-2.

Glass, G. V. and P. A. Taylor, Factor Analytic Methodology, Review of Educational Research, 1966, 36, 566-587.

Hauser, R. M., Disaggregating a social-psychological model of educational attainment in Goldberger, A. S. and Duncan, O. D. (Editors). Structural equation models in the social sciences. New York: Seminar Press, 1973.

James, W. and C. Stein, Estimation with quadratic loss. Proc. 4th Berkeley Symp. Math. and Statist. Prob., 1. Berkeley: University of California Press, 1961.

Jennrich, R. I. On the stability of rotated factor loadings: The Wexler phenomenon, Br. J. Math. Stat. Psychol., 1973, 26, 167-176.

Joreskog, K. G., On the statistical treatment of residuals in factor analysis, Psychometrika, 1962, 27, 335-354.

Joreskog, K. G. Some contributions to maximum likelihood factor analysis, Psychometrika, 1967, 32, 443-482.

Joreskog, K. G. A general method for analysis of covariance structures, Biometrika, 1970, 57, 239-257.

Lawley, D. N. A modified method of estimation in factor analysis and some large sample results, in Uppsala Symposium on Psychological Factor Analysis (Nordisk Psykolgis Monograph Series No. 3), Munsvaards, Coopenhagen; Almquist and Wiksell, Stockholm, 1953, pp. 35-42.

McDonald, R. P. Testing pattern hypotheses for covariance matrices, Psychometrika, 1974, 39, 189-201.

McDonald, R. P., Descriptive axioms for common factor theory, image theory and component theory, Psychometrika, 1975, 40, 137-152.

Mulaik, S. A. The Foundations of Factor Analysis, McGraw-Hill, New York, 1972.

Mulaik, S. A. Confirmatory Factor Analysis, in Amick, D. J. and Walberg, H.J. (Eds.), Introductory Multivariate Analysis for Educational Psychological and Social Research. Berkeley: McCutchon Publishing Corp., 1975.

Pruzek, R. M. Factor Analysis. In J. Belzer et al. (Editors) Encyclopedia of Computer Science and Technology, Vol. 8, New York: Marcel Dekker Press, 1977.

Whittle, P. A principal component and least squares method of factor analysis, Skand. Aktuarietidskift, 1952, 35, 223-239.

Table 2

A Comparison of Selected Factor Results for Hauser (1973) Correlation Matrix

Rotated PCA Factor Matrix					Rotated Whittle Factor Matrix					Rotated MLFA Factor Matrix				
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
.08	.71	.09	.41	.09	.10	.60	.09	.40	.09	.07	.10	.41	.58	.10
.11	.90	.08	.08	.07	.10	.71	.08	.15	.08	.07	.08	.16	.69	.10
.10	.26	.07	.81	.07	.12	.28	.07	.66	.08	.06	.10	.79	.23	.08
.14	.12	.05	.84	.09	.15	.18	.05	.68	.09	.04	.14	.54	.23	.15
.14	.16	.83	.12	.15	.16	.16	.66	.11	.21	.53	.20	.11	.17	.26
.19	.03	.82	-.00	.26	.21	.05	.68	.01	.29	.90	.20	.01	.05	.17
.07	.03	.25	.04	.90	.15	.05	.27	.04	.67	.33	.22	.05	.05	.45
.45	.20	.09	.18	.59	.40	.20	.15	.19	.48	.17	.29	.18	.19	.58
.69	.18	.09	.16	.19	.59	.18	.15	.19	.24	.19	.32	.20	.19	.38
.74	.13	.32	.16	.30	.60	.16	.34	.19	.34	.25	.81	.16	.16	.32
.72	.10	.34	.17	.25	.58	.13	.35	.19	.31	.28	.64	.18	.14	.35
.57	.19	.48	.21	.22	.48	.20	.44	.22	.28	.41	.44	.21	.22	.35

	1	2	3	4	5	6	7	8	9	10	11	12	Mean	S.D.	Corr.
$\hat{U}^2$ PCA	.29	.16	.26	.24	.23	.21	.10	.36	.41	.22	.26	.31	.25	.08	.133
$\hat{U}^2$ Whittle	.45	.45	.44	.46	.44	.40	.45	.49	.54	.33	.37	.39	.43	.05	.743
$\hat{U}^2$ MLFA	.46	.46	.29	.60	.56	.10	.63	.47	.63	.11	.32	.41	.42	.18	1.000

Residual Matrix $(R - \hat{R})$	PCA	Whittle	MLFA
Mean	-.0228	.00004	.00007
S.D.	.0615	.02140	.00877

	1	2	3	4	5	6	7	8	9	10	11	12
Eigenvalues (λ^2)	5.02	1.63	.864	.767	.683	.590	.530	.508	.431	.416	.342	.222
$\theta^{-2}\lambda^2$	11.56	3.76	1.99	1.76	1.57	1.36	1.22	1.17	.993	.959	.788	.512
λ^2 MLFA ($U^{-1}RU^{-1}$)	20.32	6.36	3.93	1.90	1.54	1.13	1.07	1.02	1.00	.965	.941	.868

Table 3

Comparative Statistics for Different Uniqueness Variance Estimators for Selected Data Sets

		Holzinger m=4 p=24 n=145	Holzinger m=5	Emmett m=3 p=9 n=211	Physical m=3 p=8 n=305	Hauser m=4 p=12 n=3427	Hauser m=5	Schneider m=4 p=12 n=270	Schneider m=5	Hilton m=3 p=10 n=720	Hilton m=4
	U^2 Estimators										
	MLFA	.522	.498	.372	.228	.447	.419	.287	.254	.251	.233
	MLFA/2 iter.	.523	.502	.354	.214	.456	.430	.272	.256	.251	.231
Means	Whittle	.529	.503	.372	.215	.482	.434	.272	.240	.252	.229
	$S^2(\text{Diag } R^{-1})^{-1}$	.501	.501	.478	.294	.566	.566	.347	.318	.311	.311
	PCA	.441	.398	.246	.134	.310	.253	.182	.160	.177	.138
	U^2 Estimators										
	MLFA	.151	.164	.141	.172	.172	.182	.197	.196	.071	.053
Stand.	MLFA/2 iter.	.143	.136	.129	.098	.127	.104	.148	.144	.050	.029
Dev.	Whittle	.107	.087	.085	.059	.067	.054	.086	.063	.025	.013
	$S^2(\text{Diag } R^{-1})^{-1}$	.117	.117	.147	.138	.126	.126	.192	.197	.047	.047
	PCA	.125	.095	.084	.069	.111	.083	.079	.077	.030	.018
	U^2 Estimators										
	MLFA	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Corr. with	MLFA/2 iter.	.994	.968	.910	.778	.922	.769	.869	.784	.988	.980
MLFA	Whittle	.979	.878	.779	.766	.536	.743	.863	.752	.938	.941
$\hat{U}$ values	$S^2(\text{Diag } R^{-1})^{-1}$	.919	.906	.891	.954	.752	.706	.981	.663	.931	.920
	PCA	.967	.672	.552	.143	.620	.133	.561	.396	.533	-.346
χ^2	U^2 Estimators										
	MLFA	230.80	190.64	5.82	24.21	148.62	80.74	103.85	23.10	66.68	20.15
d.f. in ()		(232)	(212)	(12)	(7)	(24)	(16)	(24)	(16)	(18)	(11)
	Whittle	329.85	288.65	68.25	190.23	1343.47	1000.81	426.03	354.80	156.09	59.05
		(209)	(189)	(20)	(14)	(35)	(27)	(35)	(27)	(27)	(20)